

# Parallelization Strategies for Spatial Agent-Based Models

Nuno Fachada<sup>1</sup>, Vitor V. Lopes<sup>2</sup>, Rui C. Martins<sup>3</sup>, and Agostinho C. Rosa<sup>1</sup>

<sup>1</sup>*Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal*

<sup>2</sup>*UTEC - Universidad de Ingeniería & Tecnología, Lima, Jr. Medrano Silva 165, Barranco, Lima, Perú*

<sup>3</sup>*INESC TEC, Campus da FEUP, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*

## Abstract

Agent-based modeling (ABM) is a bottom-up modeling approach, where each entity of the system being modeled is uniquely represented as an independent decision-making agent. Large scale emergent behavior in ABMs is population sensitive. As such, the number of agents in a simulation should be able to reflect the reality of the system being modeled, which can be in the order of millions or billions of individuals in certain domains. A natural solution to reach acceptable scalability in commodity multi-core processors consists of decomposing models such that each component can be independently processed by a different thread in a concurrent manner. In this paper we present a multithreaded Java implementation of the PPHPC ABM, with two goals in mind: 1) compare the performance of this implementation with an existing NetLogo implementation; and, 2) study how different parallelization strategies impact simulation performance on a shared memory architecture. Results show that: 1) model parallelization can yield considerable performance gains; 2) distinct parallelization strategies offer specific trade-offs in terms of performance and simulation reproducibility; and, 3) PPHPC is a valid reference model for comparing distinct implementations or parallelization strategies, from both performance and statistical accuracy perspectives.

**Keywords**— Agent-based modeling; Parallelization strategies; Shared memory; Multithreading

## 1 Introduction

Agent-based modeling (ABM) is a bottom-up modeling approach, where each entity of the system being modeled is uniquely represented as an independent

decision-making agent. When prompted to act, each agent analyzes its current situation (e.g. what resources are available, what other agents are in the vicinity), and acts accordingly, based on a set of rules (e.g. if-then-else rules, differential equations, neural networks, genetic algorithms, etc). These rules incorporate knowledge or theories about the respective low-level components. The global behavior of the system then emerges from the simple, self-organized local relationships between the agents [15]. As such, ABM is a useful tool in simulating and exploring systems that can be modeled in terms of interactions between individual agents, for example, biological cell cultures, ants foraging for food or military units in a battlefield.

Spatial agent-based models (SABMs) are a subset of ABMs in which a spatial topology defines how agents interact [51]. For example, an agent may be limited to interact with agents located within a specific radius, or may only move to a near physical or geographical location [38]. SABMs have been extensively used to study a range of phenomena in the biological and social sciences [30, 51].

Large scale emergent behavior in ABMs is population sensitive. As such, it is preferable that the number of agents in a simulation is able to reflect the reality of the system being modeled [36, 31, 29]; otherwise, the expected or desired behavior may not be observable, and model validation becomes difficult [29, 27]. This means that in domains such as social modeling, ecology, and biology, systems can contain millions or billions of individuals [12, 31, 45, 11]; consequently, simulating realistic models will involve as much agents being processed per time step [11]. Such large scale simulations generate a very high demand for computing power [26] and are impractical on typical ABM frameworks such as NetLogo [65] or Repast [43], which execute serially on the CPU [12, 11]. Additionally, stochastic models in general and ABMs in particular usually require various input parameters, which can have a range of different values. Large-scale computational experiments are required for exploring the parameter space of such models [26, 28, 56]. These requirements stretch, and many times surpass, what typical off-the-shelf computing systems can offer, especially if models are implemented in a way to only make use of one processing element (PE), such as a CPU core. Considering that commodity processors, such as GPUs and multi-core CPUs, are nowadays composed of several PEs, a natural solution to reach acceptable scalability in ABMs consists of decomposing models such that each component can be independently processed by a logical processor (LP<sup>1</sup>) in a concurrent manner [27, 55, 61, 51, 57, 9, 56]. There are, however, two main issues when parallelizing ABMs.

The first major issue concerns communication between model components and the bottleneck it creates, which is a major limiting factor in scaling parallel SABMs [57, 51]. This is especially true in distributed memory scenarios, where different computational cores may be located in separate, often geographically distant, nodes [27]. Communication costs may suppress the potential gains of using multiple nodes and their associated resources [61]. Many strategies

---

<sup>1</sup>In shared memory architectures, LPs are usually represented by threads, which communicate via synchronized access to shared variables. In distributed memory scenarios, LPs are commonly represented by processes, which communicate via message passing.

and methods have been developed to manage and reduce communication in distributed memory SABMs [50, 27], nonetheless this is still a topic of active research [51]. Whatever the scenario, model partitioning should guarantee that each model component is as independent as possible in order to minimize communication between LPs. Furthermore, communication strategies should be designed to avoid deadlocks and to preserve the causality of simulation events while efficiently exploiting parallelism [28, 51].

The second major issue when parallelizing ABMs is that it is very easy to inadvertently introduce changes which modify the model dynamic. This is akin to model replication, which is not a straightforward process [14, 67]. ABMs are very sensitive to implementation details: the impact that seemingly unimportant aspects such as data structures, algorithms, discrete time representation, floating point arithmetic or order of events can have on results is tremendous [67, 39]. The situation becomes more difficult with model parallelization, which by definition requires considerable changes in many of these aspects. Parry and Bithell [44] provide an informative account in which they were unable to successfully replicate a serial model when converting it to a parallel one. Conceptual models should be well specified and adequately described in order to achieve a successful model replication. As such, some authors have suggested that there should be a minimum standard for model communication, which should include at the very least: a) a structured natural language description using formal protocols such as ODD [25]; and, b) the model’s source code, given that it is the model’s definitive implementation, not subject to the vagueness and uncertainty possibly associated with verbal descriptions [67, 40].

In this paper we present a multithreaded Java implementation of the PPHPC<sup>2</sup> model [16], featuring several user-selectable parallelization schemes. The goals of this investigation are two-fold: 1) compare the performance of the implementation presented here, realized in a “real” programming language, with the canonical NetLogo implementation discussed in reference [16]; and, 2) study how different parallelization strategies impact simulation performance on a shared memory architecture. Care is taken so that all the parallelization strategies of the Java implementation yield similar statistical behavior as the NetLogo version, and among themselves. Leveraging on the fact that the PPHPC model captures important characteristics of SABMs, such as agent movement and local agent interactions, several conclusions on SABM parallelization techniques and the usefulness of PPHPC as a valid reference model are drawn.

The rest of the paper is organized as follows. First, in section 2, previous work about parallelization of SABMs in shared memory architectures is discussed. Next, section 3, Methodology, is composed of three parts: 1) an overview of the PPHPC model; 2) a description of the Java implementation, including the several parallelization strategies; and, 3) a discussion on how to compare multiple PPHPC versions. Results, section 4, show that: 1) model parallelization and the use of a “real” programming language (as opposed to the NetLogo modeling language) can yield considerable performance gains; 2) distinct parallelization

---

<sup>2</sup>Predator-Prey for High-Performance Computing

strategies offer specific trade-offs in terms of performance and simulation reproducibility; and, 3) PPHPC is a valid reference model for comparing distinct implementations or parallelization strategies, from both performance and statistical accuracy perspectives. Section 5 provides a global outline of what was accomplished in this work.

## 2 Background

Parry and Bithell [44] describe two techniques for partitioning SABM components across multiple computational cores: agent-parallel (AP) and environment-parallel (EP). While these are not mutually exclusive, they are nonetheless a good starting point for reasoning about SABM partitioning. In the AP approach, the model is divided at the agent-level, i.e. each LP is responsible for handling a set of agents. Load balancing is simpler as agents can be equally distributed among LPs so that each LP has a similar share of the computation [10, 44]. However, in a moving agents scenario, this partitioning leads to extra communication between LPs, which is required in order to ensure that spatially localized agent interactions are dealt with consistently, as co-location on a LP does not guarantee co-location in space [44]. In EP partitioning, model decomposition occurs at the spatial environment level, i.e. each LP is assigned a location, together with the agents it contains [10]. As such, local agent interactions will mostly occur in the same LPs. Unfortunately, when agent density varies spatially over time, e.g. in flocking or grouping patterns, load balancing issues may occur [10, 44, 22]. A corner case of this issue is when simulating chemotaxis-like patterns, where agent movement is influenced by a chemical concentration gradient, which can result in millions of agents swarming to the same location [17].

Most attempts at parallelizing ABMs found in the literature are based on the distributed memory programming model [50, 51], including a few generic ABM frameworks for high-performance computing [57, 26, 9, 7]. This approach allows models to scale to thousands of cores, usually found in supercomputer-type setups [69, 68]. However, communication issues for larger models [51] and a more complex programming paradigm (when compared with multithreading on shared memory architectures) [34] can restrict this approach.

Recently, the trend has been on hybrid [54, 2], GPU [36, 49, 63, 6] and heterogeneous [60, 62] methods. While these approaches allow concrete gains in simulation performance on commodity hardware, they come with an increased cost in implementation time due to the substantial more complex programming models. Hybrid methods, combining distributed and shared memory programming models, require modelers to master both paradigms, as well as specific multi-level model decomposition. GPU architectures require the reformulation of ABMs in terms of stream SIMD<sup>3</sup> computation and offer limited control flow constructs [36]. Heterogeneous methods, in which both CPU and GPU are utilized, entail complex synchronization and data transfers between the two

---

<sup>3</sup>Single instruction, multiple data

processors, also requiring careful model decomposition so that components can be efficiently processed.

Among the possible parallelization techniques, multithreading is arguably the simplest to implement [34, 61], with the added bonus of portability. For example, modern threading APIs, such as OpenMP<sup>4</sup> for C, C++ and Fortran or the Java 5.0 concurrency API [21], greatly simplify multithreaded ABM implementations, and are available for a number of different shared memory CPU architectures and operating systems. In this work we focus on multithreaded SABMs implemented on traditional shared-memory architectures.

While the majority of ABM toolkits [58, 46, 4, 41] are targeted for single-threaded execution on the CPU [48, 22], there have been explicit attempts to parallelize some of them. Goldsby and Pancerella [22] describe adjustments made to the MASON agent-based simulation package [35] that allow the use of multiple threads without major changes to conventional agent-based programming. RepastJ [42] has adaptations to multi-core CPUs [15], while Repast Symphony [43] supports parallel execution at the scheduling mechanism level. However, in most cases, the modeler must implement correct access semantics to shared data (e.g. environment and agent-agent interaction). One of the main problems in retrofitting parallelism to existing ABM frameworks and developing new “pure” parallel ABMs toolkits concerns the implementation of direct agent-to-agent memory access, which is model dependent and requires synchronization semantics such as locks or semaphores. These constructs are provided by the threading APIs, but efficient and thread-safe coordination of concurrent accesses still requires careful coding in order to obtain proper speedups with the number of cores, while avoiding common multithreading issues such as priority inversions or deadlocks [21].

EcoKit, a simulation system for spatially-explicit ecological models [20], was one of the first parallel realizations of an ABM on a shared memory architecture, implementing a static EP solution. The authors tested the system with a mouse migration model with 10 000 agents in a discrete 2D grid with 20 000 cells on an eight-processor SGI PowerChallenge machine [1]. One of the processors was used for the simulation kernel, while the remaining ones for the simulation itself. Results showed speedup stabilization at about four processors, reaching a maximum of 2.8 for seven processors.

A multithreaded SABM of immune system dynamics, parallelized using a static EP approach, is presented in reference [15]. Adequate speedups are obtained when chemotaxis is not simulated. However, in simulations involving this phenomena, agents in the order of thousands were shown to group in very few locations, causing severe load imbalances and limiting the scalability of simulations.

A cellular automata (CA) based ABM of opinion exchange, developed in C++ and parallelized with OpenMP, is presented by Gong et al. [23]. The authors analyze how the performance of the parallel model varies with the size of the 2D simulation grid and with the range of agent interactions. Parallelization

---

<sup>4</sup><http://openmp.org/>

of the model is achieved by partitioning the CA grid in a row-wise block-striped fashion, with each block assigned to one thread (and consequently, to one CPU core); because the agents are fixed, and there is one agent per grid cell, this type of partitioning is both AP (agent-parallel) and EP (environment-parallel). Several tests were performed in a 32-core CPU (AMD Opteron), with varying number of threads (from 1 to 32), grid sizes (up to  $5000 \times 5000$ ) and interaction ranges (from 2 to 1000 cells). Increasing the number of threads resulted in improved speedup, but lower efficiency (speedup divided by the number of threads). Lower efficiency also occurred when increasing the size of the simulation grid (maintaining the interaction range) and when increasing the interaction range. Nonetheless, the achieved speedups are impressive. For example, for a grid of  $4000 \times 4000$ , and an interaction range of 2 cells, a  $25\times$  speedup is obtained with 32 threads.

Goldsby and Pancerella [22] proposed to retrofit multithreading in the MASON ABM toolkit using an EP approach, where each thread is assigned an equal part of the simulation environment. Each simulation step is divided into two parts separated by a barrier that synchronizes access to shared environment data. Agent movement is guaranteed by interprocessor message queues. Tests were performed on a machine with 32 Intel Xeon processors using two classic ABMs: 1) Flockers, a Boids-like model [47] in which agents display flocking behavior; and, 2) the HeatBugs model [66]. The latter was shown to scale well, particularly for larger number of agents, with effective speedup up to 28 threads. The Flockers model did not scale as well, with the authors suspecting that the tendency of agents to move in flocks lead to load balancing issues.

The accurate characterization of the dynamics of bacterial networks led to the design of BNSim [64]. A multithreading approach allows BNSim to efficiently simulate large populations of bacteria. As in reference [22], simulation steps are divided in two parts with barrier synchronization, in a tick-tock pattern. During the “tick”, agents perform their actions in parallel (AP). The environment is updated in EP fashion during “tocks”. An efficient thread scheduler is used to balance the workload of both AP and EP stages of the simulation.

## 3 Methodology

### 3.1 Model overview

Here we present an overview of the PPHPC model. A complete description of the model (using the ODD protocol) is available in reference [16].

#### 3.1.1 Purpose

The purpose of PPHPC is to serve as a standard model for studying and evaluating SABM implementation strategies. It is a realization of a predator-prey dynamic system, and captures important characteristics of SABMs, such as agent movement and local agent interactions. The model can be implemented using substantially different approaches that ensure statistically equivalent results.

Entity	State variable	Symbol	Range
Agents	Type	$t$	$s, w$
	Energy	$E$	$1, 2, \dots$
	Horizontal position in grid	$X$	$0, 1, \dots, x_{\text{env}} - 1$
	Vertical position in grid	$Y$	$0, 1, \dots, y_{\text{env}} - 1$
	Energy gain from food	$g^s, g^w$	$0, 1, \dots$
	Energy loss per turn	$l^s, l^w$	$0, 1, \dots$
	Reproduction threshold	$r_T^s, r_T^w$	$1, 2, \dots$
	Reproduction probability	$r_P^s, r_P^w$	$0, 1, \dots, 100$
Grid cells	Horizontal position in grid	$x$	$0, 1, \dots, x_{\text{env}} - 1$
	Vertical position in grid	$y$	$0, 1, \dots, y_{\text{env}} - 1$
	Countdown	$C$	$0, 1, \dots, c_r$
Environment	Horizontal size	$x_{\text{env}}$	$1, 2, \dots$
	Vertical size	$y_{\text{env}}$	$1, 2, \dots$
	Restart	$c_r$	$1, 2, \dots$

**Table 1** – Model state variables by entity. Where applicable, the  $s$  and  $w$  designations correspond to prey (*sheep*) and predator (*wolf*) agent types, respectively.

Implementations may differ in aspects such as the selected system architecture, choice of programming language and/or agent-based modeling framework, parallelization strategy, random number generator, and so forth.

### 3.1.2 Entities, state variables, scales

The PPHPC model is composed of three entity classes: *agents*, *grid cells* and *environment*. Each of these entity classes is defined by a set of state variables, as shown in Table 1. Time-dependent state variables are represented with uppercase letters, while constant state variables and parameters are denoted by lowercase letters.

The  $t$  state variable defines the *agent* type, either  $s$  (*sheep*, i.e. prey) or  $w$  (*wolf*, i.e. predator). The only behavioral difference between the two types is in the feeding pattern: while prey consume passive cell-bound food, predators consume prey. Other than that, prey and predators may have different values for other state variables, as denoted by the superscripts  $s$  and  $w$ . Agents have an energy state variable,  $E$ , which increases by  $g^{\{s,w\}}$  when feeding, decreases by  $l^{\{s,w\}}$  when moving, and decreases by half when reproducing. When energy reaches zero, the agent is removed from the simulation. Agents with energy higher than  $r_T^{\{s,w\}}$  may reproduce with probability given by  $r_P^{\{s,w\}}$ . The grid position state variables,  $X$  and  $Y$ , indicate which cell the agent is located in. There is no conceptual limit on the number of agents that can exist during the course of a simulation run.

Instances of the *grid cell* entity class can be thought of the place or neighborhood where agents act, namely where they try to feed and reproduce. Agents

---

**Algorithm 1** Main simulation algorithm. **for** loops can be processed in *any order* or in *random order*. In terms of expected dynamic behavior, the former means the order is not relevant, while the latter specifies loop iterations should be explicitly shuffled.

---

```

1: INIT()
2: GETSTATS()
3:  $i \leftarrow 1$ 
4: for  $i \leq m$  do
5:   for each agent do                                     ▷ Any order
6:     MOVE()
7:   end for
8:   for each grid cell do                                   ▷ Any order
9:     GROWFOOD()
10:  end for
11:  for each agent do                                       ▷ Random order
12:    ACT()
13:  end for
14:  GETSTATS()
15:   $i \leftarrow i + 1$ 
16: end for

```

---

can only interact with other agents and resources located in the same grid cell. Grid cells have a fixed grid position,  $(x, y)$ , and contain only one resource, cell-bound food (*grass*), which can be consumed by prey, and is represented by the countdown state variable  $C$ . The  $C$  state variable specifies the number of iterations left for the cell-bound food to become available. Food becomes available when  $C = 0$ , and when a prey consumes it,  $C$  is set to  $c_r$ .

The set of all grid cells forms the *environment* entity, a toroidal square grid where the simulation takes place. The environment is defined by its size,  $(x_{\text{env}}, y_{\text{env}})$ , and by the restart parameter,  $c_r$ .

Temporal extent is represented by a positive integer  $m$ , which denotes the number of discrete simulation steps or iterations.

### 3.1.3 Process overview and scheduling

Algorithm 1 describes the simulation schedule and its associated processes. Execution starts with an initialization process, `Init()`. The process begins by instantiating the *environment* entity and filling it with  $x_{\text{env}} \times y_{\text{env}}$  grid cells. Cell-bound food is initially available with 50% probability. If not available, the countdown state variable,  $C$ , is set to a random value between 1 and  $c_r$ . At this time, a predetermined number of agents are arbitrarily placed in the simulation environment, with initial energy set to a random value between 1 and  $2g^{\{s,w\}}$ .

After initialization, and to get the simulation state at iteration zero, outputs are gathered by the `GetStats()` process (discussed in further detail in section 3.1.5). The schedule then enters the main simulation loop, where each iteration



Type	Parameter	Symbol
Size	Environment size	$x_{\text{env}}, y_{\text{env}}$
	Initial agent count	$P_0^s, P_0^w$
	Number of iterations	$m$
Dynamics	Energy gain from food	$g^s, g^w$
	Energy loss per turn	$l^s, l^w$
	Reproduction threshold	$r_T^s, r_T^w$
	Reproduction probability	$r_P^s, r_P^w$
	Cell food restart	$c_r$

**Table 2** – Size-related and dynamics-related model parameters.

is sub-divided into four steps: 1) agent movement; 2) food growth in grid cells; 3) agent actions; and, 4) gathering of simulation outputs.

In step 1, agents `Move()`, in any order, within a Von Neumann neighborhood, i.e. up, down, left, right or stay in the same cell, with equal probability. Agents lose  $l^{\{s,w\}}$  units of energy when they move, even if they stay in the same cell; if energy reaches zero, the agent dies and is removed from the simulation. In step 2, during the `GrowFood()` process, each grid cell checks if there is food available, i.e. if  $C = 0$ . If not, i.e. if  $C > 0$ ,  $C$  is decremented by one unit. In step 3, agents `Act()` in explicitly random order, i.e. the agent list should be shuffled before the agents have a chance to act. During the `Act()` process, agents first try to eat, and then try to reproduce.

### 3.1.4 Parameterization

Model parameters can be qualitatively separated into size-related and dynamics-related parameters, as shown in Table 2. Although size-related parameters also influence model dynamics, this separation is useful for parameterizing simulations.

Concerning size-related parameters, more specifically, the grid size, we propose a base value of  $100 \times 100$ , associated with 400 prey and 200 predators. Different grid sizes should have proportionally assigned agent population sizes, as shown in Table 3.

For the dynamics-related parameters, we propose two sets of parameters, Table 4, which generate two distinct dynamics. The second parameter set typically yields more than twice the number of agents than the first parameter set. Matching results with runs based on distinct parameters is necessary in order to have a high degree of confidence in the similarity of different implementations [14]. While many more combinations of parameters can be experimented, these two sets are the basis which PPHPC implementations can use to check for correctness.

While simulations of the PPHPC model are essentially non-terminating<sup>5</sup>,

<sup>5</sup>A non-terminating simulation is one for which there is no natural event to specify the

Size	$x_{\text{env}} \times y_{\text{env}}$	$P_0^s$	$P_0^w$
100	$100 \times 100$	400	200
200	$200 \times 200$	1600	800
400	$400 \times 400$	6400	3200
800	$800 \times 800$	25 600	12 800
1600	$1600 \times 1600$	102 400	51 200
3200	$3200 \times 3200$	409 600	204 800
6400	$6400 \times 6400$	1 638 400	819 200
12 800	$12 800 \times 12 800$	6 553 600	3 276 800
$\vdots$	$\vdots$	$\vdots$	$\vdots$

**Table 3** – A selection of initial model sizes. Sizes above the dashed line are studied in detail in this work.

Parameter	Symbol	Set 1	Set 2
Prey energy gain from food	$g^s$	4	30
Prey energy loss p/ turn	$l^s$	1	1
Prey reprod. threshold	$r_T^s$	2	2
Prey reprod. probability	$r_P^s$	4	10
Predator energy gain from food	$g^w$	20	10
Predator energy loss p/ turn	$l^w$	1	1
Predator reprod. threshold	$r_T^w$	2	2
Predator reprod. probability	$r_P^w$	5	5
Cell food restart	$c_r$	10	15

**Table 4** – Dynamics-related parameter sets.

the number of iterations,  $m$ , is set to 4000, as it allows to analyze steady-state behavior for all the parameter combinations discussed here.

### 3.1.5 Outputs

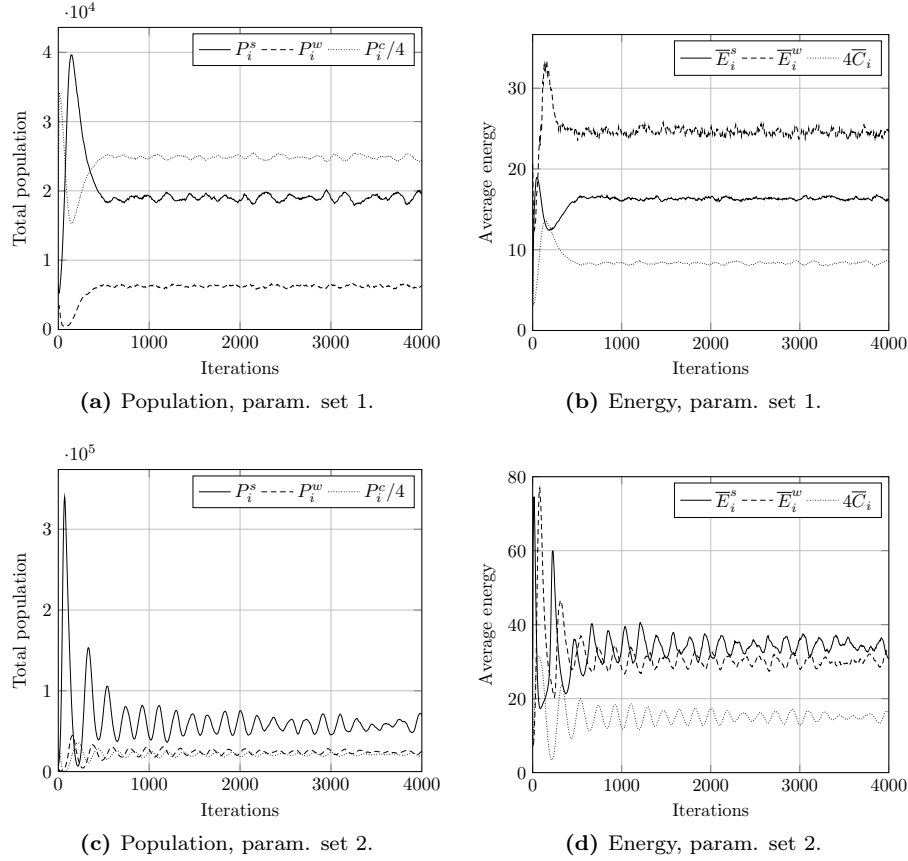
The following vector is collected in the `GetStats()` process, where  $i$  refers to the current iteration:

$$\mathbf{O}_i = (P_i^s, P_i^w, P_i^c, \bar{E}_i^s, \bar{E}_i^w, \bar{C}_i)$$

$P_i^s$  and  $P_i^w$  refer to the total prey and predator population counts, respectively, while  $P_i^c$  holds the quantity of available cell-bound food.  $\bar{E}_i^s$  and  $\bar{E}_i^w$  contain the mean energy of prey and predator populations. Finally,  $\bar{C}_i$  refers to the mean value of the  $C$  state variable in all grid cells.

---

length of a run [33].



**Figure 1** – Typical model output for model size 400. Other model sizes have similar outputs, apart from a vertical scaling factor.  $P_i$  refers to total population,  $\bar{E}_i$  to mean energy and  $\bar{C}_i$  to mean value of the countdown state variable,  $C$ . Superscript  $s$  relates to prey,  $w$  to predators, and  $c$  to cell-bound food.  $P_i^c$  and  $\bar{C}_i$  are scaled for presentation purposes.

Figure 1 shows the typical output of a simulation run with size 400 for both parameter sets. Model sizes up to 12 800 have similar outputs, apart from a vertical scaling factor.

### 3.2 A multithreaded Java implementation

Java is a general-purpose, object-oriented computer programming language, and is usually compiled to bytecode that runs on a Java Virtual Machine (JVM) [24]. Version 5.0 of the Java programming language marked a huge step forward for the development of concurrent applications, offering new higher-level components and additional low-level mechanisms that make it simpler to build concurrent applications [21]. Java is a well-known language within the ABM



---

**Algorithm 2** Realization of the main simulation algorithm in the `SimWorker` class. Calls to `ControllerSync()` are controller synchronization points.

---

```

1: CONTROLLERSYNC(1)
2: CREATECELLS()                                ▷ Environment-parallel
3: CONTROLLERSYNC(2)
4: SETCELLNEIGHBORS()                            ▷ Environment-parallel
5: CONTROLLERSYNC(3)
6: CREATEPREY()                                ▷ Agent-parallel
7: CREATEPREDATORS()                            ▷ Agent-parallel
8: CONTROLLERSYNC(4)
9: GETSTATS()                                    ▷ Environment-parallel
10: CONTROLLERSYNC(5)
11:  $i \leftarrow 1$ 
12: for  $i \leq m$  do
13:   MOVE() + GROWFOOD()                        ▷ Environment-parallel
14:   CONTROLLERSYNC(6)
15:   ACT() + GETSTATS()                         ▷ Environment-parallel
16:   CONTROLLERSYNC(7)
17:    $i \leftarrow i + 1$ 
18: end for
19: CONTROLLERSYNC(8)

```

---

1 in a different order, but in a way that the final qualitative simulation outcome does not change, as outlined in Algorithm 2<sup>6</sup>.

The `Init()` process, line 1 of Algorithm 1, is divided in four steps, which can be processed in parallel, as determined by the work provider: 1) instantiation and initialization of grid cells (line 2 of Algorithm 2); 2) connecting cells to their neighbors (line 4 of Algorithm 2); 3) instantiation of prey (line 6 of Algorithm 2); and, 4) instantiation of predators (line 7 of Algorithm 2). In steps 1 and 2 the unit of work represents a single grid cell. In steps 3 and 4 the unit of work represents the instantiation and deployment of a single agent.

Next, workers execute the `GetStats()` process (line 2 of Algorithm 1 and line 9 of Algorithm 2). Data required for the observation of the model, specified in section 3.1.5, is collected on a cell-by-cell basis. After processing their allocated cells, workers then update a global statistics object.

Lines 5 to 10 of Algorithm 1, which include agent movement and growth of cell-bound food, are condensed into a single EP loop (line 13 of Algorithm 2). This is possible because the `Move()` and `GrowFood()` processes are independent; i.e. the consequences of either will only impact the `Act()` process, which occurs later. Most importantly, both `Move()` and `GrowFood()` are cell-wise independent and can be processed autonomously for each cell in an EP loop. When a cell is processed, agents located therein are prompted to move, and then the cell

---

<sup>6</sup>Controller synchronization points (i.e. calls to `ControllerSync()`) are discussed in section 3.2.3.

is asked to execute its `GrowFood()` process. Care is taken so that agents that already moved are not prompted to move again.

Finally, lines 11 to 14 of Algorithm 1, containing the `Act()` and `GetStats()` processes, are also contracted into one EP loop, as shown in line 15 of Algorithm 2. While agent actions and statistics gathering are not independent events (i.e. the former must occur before end-of-iteration data is obtained), they are cell-wise independent. As specified in section 3.1.2, the actions of an agent are limited to the cell it occupies. Thus, after agent actions take place in a cell, end-of-iteration cell data is collected. Note that before agents in a cell act, the local agent list is shuffled according to Algorithm 1, i.e. *random order*. After processing all of their allocated cells, workers then update a global statistics object.

Lines 13 and 15 of Algorithm 2 reflect the separation of each simulation step into two parts, in a “tick-tock” fashion. This is often necessary in multi-threaded ABM implementations to synchronize agent movement, agent actions, environmental dynamics or reading and writing of shared environment areas [64, 22].

### 3.2.2 Random number generators and reproducible simulations

Reproducibility of simulations is an important and often overlooked aspect of ABM parallelization. As explained by Hill et al. [28], “To investigate and understand the results, we have to reproduce the same scenarios and find the same confidence intervals every time we run the same stochastic experiment. When debugging parallel stochastic applications, we need to reproduce the same control flow and the same result to correct an anomalous behavior”. Most ABMs, including the one presented here, use one or more stochastic processes.

PRNGs use iterative deterministic algorithms for producing a sequence of pseudo-random numbers that approximate a truly random sequence [8]. A PRNG consists of a finite set of states, and a transition function that takes the PRNG from one state to the next. The initial state of the PRNG is called the seed [52]. As such, PRNGs are used in simulations to mimic stochastic processes in a reproducible fashion.

Reproducibility is simple to accomplish in a single-threaded implementation, sufficing the use of the same PRNG and seed, with deterministic scheduling of all stochastic processes. However, in a parallel simulation context, there are practical trade-offs between reproducibility, memory and speed [61]. Additionally, the use of PRNGs in parallel simulations comes with its own set of problems, such as hidden correlations or overlaps in different sub-streams of the same PRNG [28].

For the Java PPHPC implementation, parallel pseudo-random generation is carried out in the following manner. The  $i^{\text{th}}$  worker thread obtains its own sub-sequence of a global random sequence by using a unique seed,  $S_i$ , through a random spacing approach [28]. Each seed  $S_i$  is derived from the worker’s unique identifier,  $i$ , and from a user specified global seed,  $S$ , according to Eq. 1,

$$S_i = \begin{cases} S, & \text{if } i = 0 \\ S \oplus \text{SHA256}(i), & \text{if } i > 0 \end{cases} \quad (1)$$

where  $\oplus$  is the bitwise XOR operator and  $\text{SHA256}()$  is the SHA-256 cryptographic hash function. The Java implementation of PPHPC uses the Uncommons Maths library [13], taking advantage of the several PRNGs it provides. For the results presented in this paper, the library’s Mersenne Twister implementation was used, because it is the same PRNG used by NetLogo, and its very large period of  $2^{19937} - 1$  makes sub-stream overlapping highly unlikely to occur [5, 53].

Considering that each worker has its own PRNG sub-sequence, the following conditions are required in order to make simulations reproducible:

1. Each worker must process the exact same work between runs, i.e. it must:
  - (a) Instantiate the same quantity of initial agents and place them in the same cells.
  - (b) Process the same cells.
2. Agents within a cell must be processed in the same order.

The first condition can be guaranteed if work providers always assign the same tokens to the same worker. The second condition may be problematic when cell-level synchronization is required, which occurs during agent movement. As will be discussed in the next section, this issue can be solved within the current framework by placing agents in their destination cell using some deterministic order criteria. This ensures that, when entering the agent action stage, each cell contains an ordered list of agents, ready to act.

A broader way of forcing reproducibility of simulations is to associate the PRNG sub-sequences with cells and not with worker threads. However, this can result in a large number of parallel PRNG states (e.g.  $1.6 \times 10^5$  PRNG states are required for model size 400). There are two main problems with this approach. The first, already briefly discussed, is that the problems associated with parallel PRNGs become worse when partitioning the PRNG stream into more and more sub-streams [28]. The second issue is related to the required memory. For the problem of size 400, using a large period PRNG (in order to minimize the first problem) such as the Mersenne Twister, will approximately require 380 MiB of memory just for the PRNG states; for larger model sizes, such as 1600, almost 6 GiB are required.

### 3.2.3 Parallelization strategies

A parallelization strategy is defined by the selected work factory, more specifically by the work providers it offers, and by the way it configures the controller. The work factory is first requested to instantiate and configure an appropriate controller object. When the simulation starts, the controller spawns the number

of workers specified by the work factory, and each worker gets a reference to the controller and to the work factory. Workers use the former to synchronize themselves and the latter to get references to work providers; these, in turn, provide workers with tokens, i.e. integers that uniquely identify units of work. When a work provider returns a negative value, it means that the current parallel work cycle is finished. Three work providers are used by workers: one for the cells, which provides work in an EP fashion for all cell-wise work cycles, and two for the initial agent creation (one for prey, the other for predators), which provide AP work. The latter are used only once during the `Init()` process, while the former, i.e. the cell work provider, is continuously reused.

To accommodate different parallelization strategies, workers have several possible synchronization points, which occur at three different levels: 1) controller; 2) work provider; and 3) grid cells.

There are eight controller-level synchronization points. All workers explicitly notify the controller when they reach them, as shown by the calls to `ControllerSync()` in Algorithm 2. Whether or not workers are held on that point by the remaining workers depends on how the controller was configured by the work factory. There are, however, some points where barriers are mandatory. For example, no worker can begin processing agent actions before all agents have moved and all food has grown; thus, synchronization point 6 (line 14 of Algorithm 2) is necessarily a barrier. Sequences of concurrent computation (AP or EP) and thread communication, terminating with a barrier, can be considered global *supersteps* under a Bulk Synchronous Parallel model [59].

Work provider and cell-level synchronization is performed implicitly when workers request work and when they process cells, respectively. There are two possible cell-level synchronization points: 1) when inserting initial agents during the `Init()` process; and, 2) when inserting agents moving from other cells. Again, whether or not workers are actually synchronized at work provider and cell-level synchronization points depends on the parallelization strategy. Five parallelization strategies are provided, as shown in Table 5, which also enumerates all possible synchronization points and how the different strategies handle them. The following paragraphs describe in detail these parallelization strategies.

**Single-thread (ST)** A single-threaded work factory (`SingleThreadWorkFactory` class) is provided for comparison with the multithreaded work factories. This work factory configures the controller such that no synchronization occurs when the (single) worker explicitly notifies the controller that it reached a given synchronization point. Likewise, no work provider or cell-level synchronization is required. The work providers made available by the single-threaded work factory (instances of the `SingleThreadWorkProvider` class) maintain a simple counter which issues tokens to the single worker. Simulation objects (cells and agents) are always iterated in the same order, which allows for simulation reproducibility.



PS	Controller								WP	Cell	
	1	2	3	4	5	6	7	8		Init.	Move
ST	-	-	-	-	-	-	-	-	-	-	-
EQ	$S$	$B$	$S$	$B$	$S$	$B$	$B$	$S$	-	$S$	$S$
EX	$S$	$B$	$S$	$B$	$S$	$B$	$B$	$S$	-	$\hat{S}$	$\hat{S}$
ER	$S$	$B$	$S$	$B$	$S$	$B$	$B$	$S$	$B$	$\hat{S}$	-
OD	$S$	$B$	$S$	$B$	$B$	$B$	$B$	$S$	$S$	$S$	$S$

**Table 5** – Parallelization strategies (PS) and their handling of the possible synchronization points at the level of the controller, work provider (WP) and grid cell.  $B$  means there is a barrier, i.e. that workers can only advance when all workers have reached the sync. point.  $S$  implies access serialization, but workers do not have to wait on other workers before continuing;  $\hat{S}$  is similar, but implies an ordered agent insertion, which might take longer than simple access serialization.

**Equal (EQ)** The general idea of the EQ parallelization strategy (handled by the `EqualWorkFactory` class) is that each worker always processes the same work. Work distribution is performed once at the beginning of the simulation by the associated work providers (instances of the `EqualWorkProvider` class), and then the workers are always given the same exact tokens, e.g. they always process the same cells in the EP sections. Cell-level synchronization is required because more than one worker may potentially access the same cell at the same time for agent movement or initial agent placement. The first worker to whom access is granted gets to place its agent topmost in the cell’s internal agent list. This means that, because thread synchronization is not a deterministic process, agents will not be processed in the same order from run to run. Thus, simulations with this work factory are not reproducible. The maximum number of tokens to be processed by each worker,  $n$ , is given by

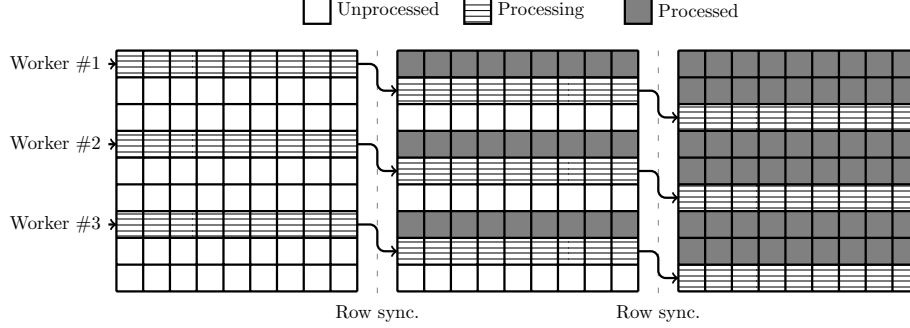
$$n = \lceil T/N \rceil \quad (2)$$

where  $T$  is the number of tokens to be processed in a parallel work cycle, and  $N$  is the number of worker threads. If  $T$  is not equally divisible between the available workers, the last worker will process less work than the remaining workers, as shown in Eq. 3,

$$n \cdot i \leq t_i < \min(n \cdot (i + 1), T) \quad (3)$$

where  $i$  identifies the  $i^{\text{th}}$  worker, and  $t_i$  corresponds to the range of tokens which will be processed by the  $i^{\text{th}}$  worker.

**Equal with repeatability (EX)** The EX parallelization strategy is a slight variation on EQ. The same classes are used, the only difference is that when workers access the same cell at the same time for agent placement purposes,



**Figure 3** – Equal with row synchronization (ER): example of three workers processing nine rows of the simulation grid in parallel.

the agent is placed in ordered fashion, as shown in Table 5. This allows for reproducible simulations because agents will be processed in the same order.

**Equal with row synchronization (ER)** In the ER parallelization strategy, as in the previous strategies, work is assigned to workers at the beginning of the simulation. The difference here is that each worker serially processes rows of the simulation grid, leaving a distance of at least three rows (including the row to be processed) to the next worker (see Figure 3). More generally, this distance is given by

$$d_{\min} = 2r + 1 \quad (4)$$

where  $r$  is the agent movement radius, which is 1 for the PPHPC model. This approach allows workers to run in parallel without any need for cell-level synchronization, because they synchronize at the end of each row at the work provider level. Thus, agents always move to neighboring cells in the same order, making simulations reproducible.

The `EqualRowSyncWorkFactory` class is responsible for configuring the controller and issuing the appropriate work providers. An instance of the `EqualRowSyncWorkProvider` class is issued for EP work. However, this approach does not make sense for AP work; as such, the agent initialization phase, which is handled in AP fashion, is managed by instances of the `EqualWorkProvider` class. For simulations to be reproducible, initial agents are inserted in an ordered fashion, as shown in Table 5.

This strategy implies that there is a practical maximum number of workers depending on the number of rows,  $y_{\text{env}}$ , and on the minimum distance between rows,  $d_{\min}$ , as shown in Eq. 5:

$$N_{\max} = \left\lfloor \frac{y_{\text{env}}}{d_{\min}} \right\rfloor \quad (5)$$

If the specified number of workers,  $N$ , is larger than  $N_{\max}$ , an exception is

thrown and the simulation terminates. An initial estimate of the number of rows per worker,  $\Delta y$ , is given by

$$\Delta y = \left\lfloor \frac{y_{\text{env}}}{N} \right\rfloor \quad (6)$$

This estimate can be incremented if: 1) the number of rows is not equally divisible by the number of workers; and, 2) after incrementing it, there are enough rows for the last worker to process. This is shown in Eq. 7, where  $\Delta y_f$  is the final number of rows per worker:

$$\Delta y_f = \begin{cases} \Delta y + 1, & \text{if } y_{\text{env}} \bmod N > 0 \wedge (N - 1) \cdot (\Delta y + 1) \leq y_{\text{env}} - d_{\min} \\ \Delta y, & \text{otherwise.} \end{cases} \quad (7)$$

From the workers perspective, what matters are the tokens to process. All workers, except possibly the last one, will process  $n$  tokens, according to Eq. 8. The exact tokens that the  $i^{\text{th}}$  worker will process are given in Eq. 9. Note that any adjustment due to the number of rows not being exactly divisible by the number of workers is performed on the last worker.

$$n = x_{\text{env}} \cdot \Delta y_f \quad (8)$$

$$n \cdot i \leq t_i < t_f, \text{ where } t_f = \begin{cases} n \cdot (i + 1), & \text{if } i < N - 1 \\ x_{\text{env}} \times y_{\text{env}}, & \text{if } i = N - 1 \end{cases} \quad (9)$$

**On-demand (OD)** The OD parallelization strategy, managed by an instance of the `OnDemandWorkFactory` class, aims to improve load balancing by issuing smaller blocks of tokens to keep workers busy. Work providers, instances of the `OnDemandWorkProvider` class, maintain a counter of the tokens already issued to workers. Each time a worker requests more tokens, this counter is incremented by the block size,  $b$ . Access to the work provider is serialized, as shown in Table 5, because the work counter needs to be atomically incremented. The counter is implemented using an instance of the `AtomicInteger` class, which was added to the Java SE 5.0 API. Lower values of  $b$  will cause workers to fetch tokens from the work provider more often, which may cause some thread contention; however, work distribution is improved because workers are more likely to be processing work instead of waiting for slower workers at controller synchronization points. Conversely, with higher values of  $b$ , worker threads will request work less frequently, which leads to lower thread contention; on the downside, faster workers may have to wait longer for slower ones. Thus,  $b$  controls a trade-off between thread contention and load balancing. If  $b$  is selected such that workers only access the work provider once and the same amount of work is assigned to each worker, the OD strategy becomes similar to EQ, but with additional synchronization.

Table 5 also shows that the OD strategy requires all workers to explicitly wait for one another at controller synchronization point 5. This is required

because some workers may finish processing their `GetStats()` tokens early (line 9 of Algorithm 2), while others lag behind; at line 13 of Algorithm 2, faster workers could potentially obtain work tokens that are still being processed by slower workers at line 9.

This parallelization strategy does not offer reproducible simulations because workers obtain tokens in a FIFO fashion which is dependent on the OS thread scheduling, and thus not deterministic. As such, it is not possible to anticipate which worker will process which tokens, resulting in cells being associated with different worker-bound PRNG sub-sequences from iteration to iteration and from run to run.

### 3.2.4 Source code

The Java implementation described here is available at <https://github.com/fakenmc/pphpc/tree/java/java>.

## 3.3 Comparison of multiple implementations

Axtell et al. [3] define three replication or comparison standards for the level of similarity between model outputs: *numerical identity*, *distributional equivalence* and *relational alignment*. The first, *numerical identity*, implies exact numerical output, but it is difficult to attain and not critical for showing that two models have the same type of dynamic behavior. To achieve this goal, *distributional equivalence* is a more appropriate choice when parallelizing ABMs, as it aims to reveal the statistical similarity between two outputs. Finally, *relational alignment* between two outputs exists if they show qualitatively similar dependencies with input data, which is frequently the only way to compare a model with another which is inaccessible (e.g. implementation has not been made available by the original author), or with a non-controllable “real” system (such as a model of the human immune system [15]).

Given that we are essentially interested in showing that the parallel variants of the Java implementation have statistically indistinguishable behavior from the canonical NetLogo implementation, we aim for the comparison standard of distributional equivalence. In order to compare multiple versions of PPHPC for distributional equivalence, we follow the approach taken by Wilensky and Rand [67], demonstrating the distributional equivalence of a few statistical summaries for each output, avoiding the need of trying to show distributional equivalence for every output at every iteration. This process can be realized with the following steps: 1) select a set of statistics which summarize each output (individual statistics from each output are designated as focal measures); 2) for each version of PPHPC or parallelization strategy, obtain samples of individual focal measures from a number of replications (each replication performed with a different PRNG seed); and, 3) apply a statistical test to the several samples of individual focal measures to check if they are distributionally equivalent.

In the first step we select, for each of the six outputs, the six statistical summaries described in Table 6, in a total of 36 focal measures. The rationale

Statistic	Description
$\max_{0 \leq i \leq m} X_i$	Maximum value.
$\arg \max_{0 \leq i \leq m} X_i$	Iteration where maximum value occurs.
$\min_{0 \leq i \leq m} X_i$	Minimum value.
$\arg \min_{0 \leq i \leq m} X_i$	Iteration where minimum value occurs.
$\bar{X}^{ss} = \sum_{i=l+1}^m X_i / (m - l)$	Steady-state mean.
$S^{ss} = \sqrt{\frac{\sum_{i=l+1}^m (X_i - \bar{X}^{ss})^2}{m - l - 1}}$	Steady-state sample standard deviation.

**Table 6** – Statistical summaries for each output  $X$ , where  $X_i$  is the value of  $X$  at iteration  $i$ ,  $m$  denotes the last iteration, and  $l$  corresponds to the iteration separating the transient and steady-state stages. For parameter set 1,  $l = 1000$ , while for parameter set 2,  $l = 2000$ .

behind this selection, including the choice of  $l$ , which separates the transient and steady-state stages of the simulation, is discussed in detail in reference [16].

For step 2, a number of replications was performed for each version of PPHPC or parallelization strategy, with the global PRNG seed,  $S$ , taken from the MD5 checksum of the replication number such that individual replications are performed with essentially uncorrelated seeds.

Concerning step 3, a number of statistical tests can be used to determine if focal measures from different PPHPC versions or parallelization strategies are distributionally equivalent. The choice of statistical test(s) depends on a number of factors, such as the number of samples to be compared or whether a specific focal measure follows a normal distribution. Regarding the former, there are six samples to be compared for each focal measure, produced by the NetLogo implementation and the five parallel variants of the Java implementation. Concerning the latter, some focal measures follow an approximately normal distribution, while others do not [16]. Consequently, a statistical test which is able to simultaneously compare six samples and does not assume normality of sample populations is desirable. The Kruskal-Wallis test [32] fills these requirements. It is the non-parametric equivalent of ANOVA, allowing us to compare samples drawn from populations with normal and non-normal distributions, extending the Mann-Whitney U test for more than two samples [19]. It tests the null hypothesis that all samples are drawn from the same distribution by comparing their medians. As such, it is used in this work to compare the outputs of the six PPHPC realizations.

## 4 Results and Discussion

A total of 10 replications were performed with the following combination of parameters:

- *Implementations (variants)*: NetLogo (NL) and Java (ST, EQ, EX, ER and OD)
- *Parameter sets*: 1, 2
- *Model sizes*: 100, 200, 400, 800, 1600
- *Number of threads (multithreaded Java variants only)*: 1, 2, 4, 6, 8, 12, 16, 24
- *Block size (Java OD variant only)*: 20, 50, 100, 200, 500, 1000, 2000, 5000

This implies a total of 20 and 1780 replications per size for the NetLogo and Java implementations, respectively, causing practical difficulties in scaling the experiment to sizes above 1600. For each combination of parameters, the 10 replications were performed with distinct PRNG seeds. These replications are the basis for both the performance and statistical analysis performed in the next sections. All performance results are based on the mean run time of the 10 replications.

Replications were performed in “headless” mode for both implementations, i.e., without any graphical component. For the NetLogo implementation this meant the use of the BehaviorSpace tool from the command line. Parallel runs were disabled because we were interested in benchmarking the performance of individual runs, and simultaneous runs may have interfered in the measurements. Replications with the Java implementation were performed with a single non-GUI MVC view, which performs a simulation from start to finish without user intervention.

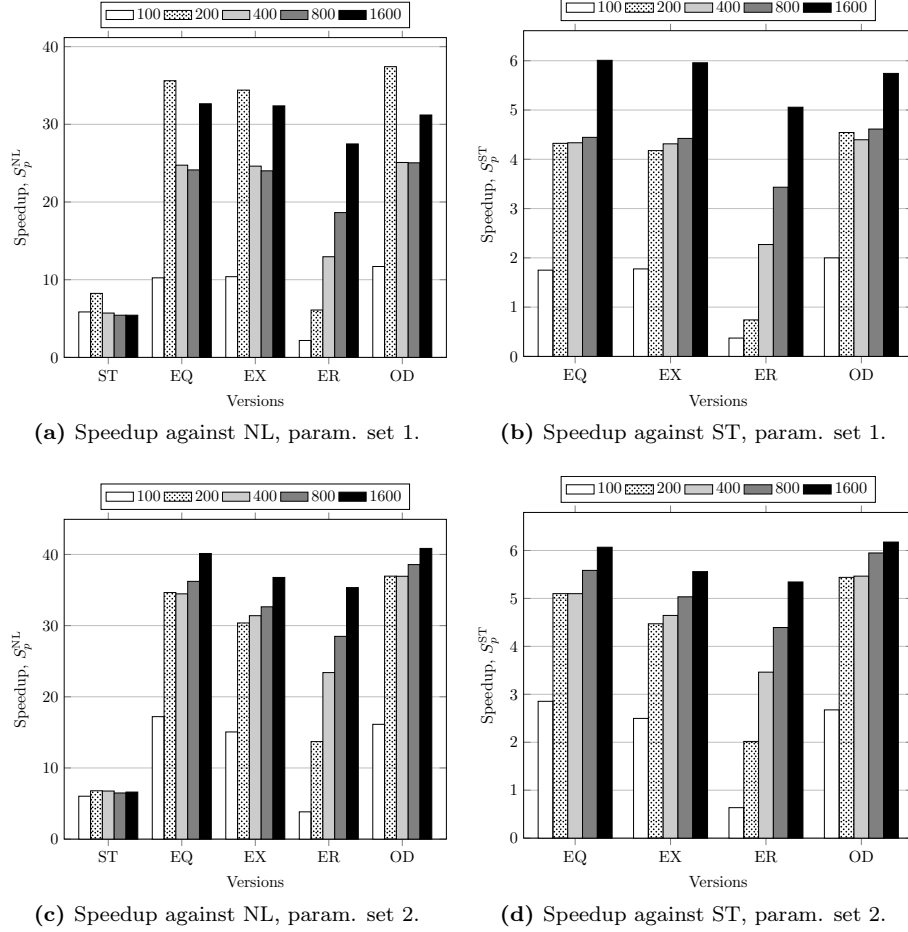
All replications were performed on a machine with the following hardware and software configuration:

- Intel(R) Core(TM) i7-3930K CPU 3.20GHz (six cores, two logical processors per core), 32GB RAM
- Ubuntu 14.04.2 LTS, OpenJDK Java 1.7.0, NetLogo 5.1.0

The data produced by this computational experiment, as well as the scripts used to set up the experiment, are made available to other researchers at <https://zenodo.org/record/34049>.

### 4.1 Performance comparison

Figure 4 shows, for both parameter sets, the speedups of the several parallel variants against the NetLogo and Java single-thread versions. Speedup is the execution time of the single-threaded versions, either  $T_1^{\text{NL}}$  for NetLogo, or  $T_1^{\text{ST}}$

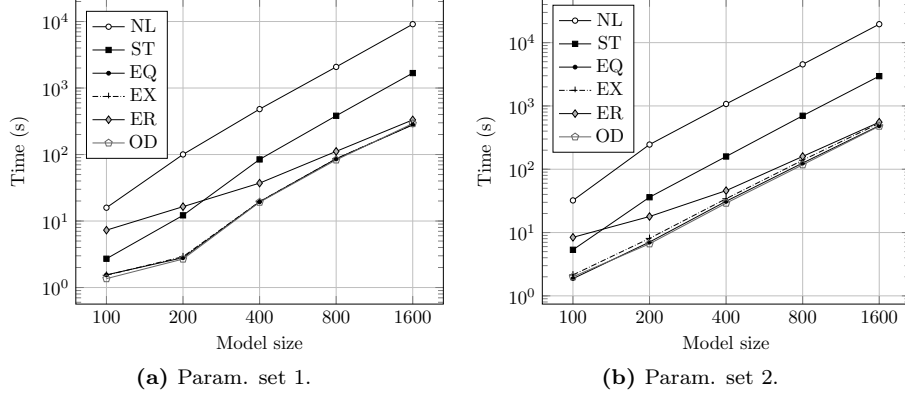


**Figure 4** – Speedup for parallel variants with 12 workers against single-thread versions, with  $b = 500$  for the OD variant.

for Java single-thread, over the execution time of a parallel variant using  $p$  workers ( $T_p$ ), as defined by

$$S_p^{\{NL,ST\}} = \frac{T_1^{\{NL,ST\}}}{T_p} \quad (10)$$

Figure 5 establishes how the different versions scale with increasingly larger model sizes. The results shown in Figures 4 and 5 were obtained with 12 worker threads for the Java multithreaded variants, and with  $b = 500$  for the OD variant. The processor used can concurrently handle 12 threads. With the exception of model size 100, we find that  $N = 12$  is the number of workers that offers the best performance. Concerning the OD variant,  $b = 500$  consistently



**Figure 5** – Scalability of the different versions for increasing model sizes. Parallel variants are using 12 workers, with  $b = 500$  for the OD variant.

provides good performance across the different model sizes and parameter sets. Table 7 provides more detailed data, including simulation times and relative standard deviations. The latter were consistently small, mainly in the order of 1%-4%, and are thus not considered for the remainder of the discussion.

In Figure 4 it is immediately noticeable that the single-threaded version of the Java implementation (ST) is  $5\times$  to  $8\times$  faster than the NetLogo (NL) implementation. Interestingly, the speedup is very consistent for the several tested sizes and for both parameter sets. Although NetLogo is getting considerably more efficient in the last few years [37], implementing an ABM directly in a programming language still seems to provide better performance. Observing the results for the multithreaded Java variants (EQ, EX, ER and OD), speedups up to 40 further validate that hypothesis.

The Java implementations with equal work distribution, EQ and EX, offer similar performance, with a slight advantage to the former. The ordered agent insertion performed by the EX variant takes a small toll, but the advantage of repeatability clearly offsets that issue.

It is also obvious that the potential benefits offered by the ER variant, namely the avoidance of cell-level synchronization during agent movement, do not outweigh the overhead of row-level synchronization. However, as shown in Figures 4 and 5, the performance of ER improves for larger model sizes. Nonetheless, the EX variant also offers repeatability, while being faster for the tested model sizes. The ER variant is actually slower than the ST version for smaller simulations. The only selling point for ER would be if it could outperform other implementations for very large model sizes not tested in this work. This is suggested by Figure 5, which shows that ER is the only variant that decreases relative simulation time with increasing model sizes.

The OD variant offers the best speedups in most of the test cases. The problem with the OD strategy is that it does not offer repeatable simulations,



although a more general solution for this problem is discussed in section 3.2.2. A performance analysis specific for the OD version is presented in section 4.3.

Generally, as shown in Figure 4, the speedup of the multithreaded variants improves for larger model sizes. This is clearer for parameter set 2, which yields simulations with more agents. Larger models, in grid size or number of agents, tend to reduce the overhead of parallelization [22]. The only exception is for model size 200 with parameter set 1, where all pure Java versions offer substantial speedups versus the NL version (see Figure 4a). Figure 5a shows this variation can be attributed to consistently faster than expected execution of the Java implementations (except ER) for this size. Although the variation is unexpected, the individual run times do not show much variance between themselves (as shown in Table 7), so this behavior is consistent and not attributable to outliers. With model size 1600, most parallel variants offer speedups up to 6 against the ST version, denoting efficient usage of the six-core hyper-threaded Intel processor. If we consider the number of LPs as the performance target, however, near ideal speedup,  $S_p^{ST} \approx 12$ , is harder to achieve because each pair of LPs shares execution resources from a single hyper-threaded core.

## 4.2 Varying the number of worker threads

Figure 6 displays, for the two parameter sets and selected model sizes (100, 400 and 1600), how the simulation time is affected by varying the number of worker threads in the multithreaded variants. Except for model size 100, the optimal number of workers is 12, which is in accordance with the number of threads directly supported by the hardware. As a general tendency, larger model sizes benefit more from higher number of workers. The same can be said of parameter set 2, which, having considerably more agents in action, also profits from having more workers. For smaller sizes and/or lower number of agents, each worker thread has less work to do and thus workers spend a larger percentage of time waiting in synchronization points. The OD variant seems to scale better with more workers, most likely because workers process work as it becomes available, spending relatively less time waiting.

## 4.3 Analysis of the OD parallelization strategy

The OD parallelization strategy deserves a more thorough analysis because of the block size parameter,  $b$ . Figure 7 shows, for the two parameter sets and different model sizes, how simulation times vary with several values of  $b$ . The number of workers is set to 12. The values of  $b$  that offer faster simulation times are marked with a bold circle. Although there is a slight tendency for higher values of  $b$  to be associated with larger model sizes, what stands out from Figure 7 is that performance is not much affected by  $b$ , with the exception of size 100; in this case, larger values of  $b$  divide the environment in sizeable sections, such that some workers never get any work to do. For example, a  $100 \times 100$  grid has 10 000 grid cells, and if divided into sections of 5000 cells, only two worker threads will have work to do.

Generally, best or close to best simulation times are obtained with  $b = 500$  across all tested model sizes and parameter sets. As such, this is the value used when comparing OD with the remaining versions.

#### 4.4 Statistical analysis of model output data

The  $p$ -values obtained from the Kruskal-Wallis test by comparing the selected focal measures of the different PPHPC versions for all combinations of parameter sets and model sizes are provided in Table 8. In a total of 360  $p$ -values (36 focal measures, 2 parameter sets, 5 model sizes), 28 are below the 0.05 significance level; of these, only nine are below 0.01. However, there are two low  $p$ -values that stand-out for  $\min \bar{E}_i^s$ , associated with model sizes 200 and 800, parameter set 1, suggesting that at least one of the implementations (i.e. samples) produced significantly different minimum values for mean prey energy. Taking a closer look at the distributional behavior of the  $\min \bar{E}_i^s$  focal measure for parameter set 1 [16], we note that this value always occurs at iteration zero, the initial state of the simulation (i.e.,  $\arg \min \bar{E}_i^s = 0$ ). Additionally, as described in section 3.2, the EQ and EX parallelization strategies share the same work provider for agent initialization, which means, that for a given PRNG seed, both strategies generate initial agents with the exact same energy. As such, if the initial prey energy is somewhat different from usual for one of these strategies, it will be so for both. Consequently, the Kruskal-Wallis test compares two samples with somewhat different observations (EQ and EX) from the remaining four (NL, ST, ER and OD), which is the case for sizes 200 and 800. If we remove EQ and EX from the test, the  $p$ -values rise to 0.128 and 0.002 for sizes 200 and 800, respectively, in line with the remaining results. Other than this, the  $p$ -values do not appear to follow any trend or pattern, e.g. smaller  $p$ -values do not seem to be associated with any particular focal measure, parameter set or model size.

From these results, it is possible to conclude that all realizations appear to produce similar dynamic behavior. Thus, model parallelization does not seem to have introduced any observable bias for the tested parameter sets.

## 5 Conclusions

In this paper, a Java implementation of the PPHPC model, providing several multithreaded parallelization strategies, was proposed. The model captures important characteristics of SABMs, such as agent movement and local agent interactions. Three conclusions are drawn from this study:

1. SABM parallelization, if done carefully, can yield considerable performance gains, with speedups up to 40 on a six-core hyper-threaded processor, and maintain statistical accuracy with the original serial model. While developing models in NetLogo is much simpler than directly using Java or other programming languages, it comes at a considerable cost in terms of performance.

2. Different parallelization strategies offer specific trade-offs in terms of performance and simulation reproducibility. For example, the generally most efficient parallelization strategy, OD, does not allow for reproducible simulations and exposes additional complexity due to the additional block size parameter. The EX variant, however, does support reproducible simulations and is not far off in terms of performance.
3. PPHPC was shown to be a valid reference model for comparing distinct implementations or parallelization strategies, from both performance and statistical accuracy perspectives.

The Java implementation of PPHPC and the respective parallelization strategies can be subject to additional experimentation, e.g. in machines with higher core counts or by exploring additional parameter sets. Nonetheless, future implementations in distinct architectures (e.g. GPU, FPGA, distributed memory) or programming languages, should provide an opportunity to further evaluate how can SABMs be appropriately implemented in each case.

## 6 Acknowledgments

This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) projects UID/EEA/50009/2013, UID/MAT/04561/2013 and (P. RD0389) Incentivo/EEI/LA0009/2014, and partially funded with grant SFRH/BD/48310/2008, also from FCT. The author Vitor V. Lopes acknowledges the financial support from the Prometeo project of SENESCYT (Ecuador).

## References

- [1] The POWER CHALLENGE technical report. Technical report, Silicon Graphics, Inc., 1994.
- [2] B. Aaby, K. Perumalla, and S. Seal. Efficient simulation of agent-based models on multi-GPU and multi-core clusters. In *Proceedings of the 3rd International ICST Conference on Simulation Tools and Techniques*, pages 1–10. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2010.
- [3] R. Axtell, R. Axelrod, J. M. Epstein, and M. D. Cohen. Aligning simulation models: a case study and results. *Comput. Math. Organ. Theory*, 1(2):123–141, 1996.
- [4] M. Berryman. Review of software platforms for agent based models. Technical Report DSTO-GD-0532, Land Operations Division, Defence Science and Technology Organisation, PO Box 1500, Edinburgh, South Australia 5111, Australia, Apr. 2008.

- [5] C. Brugger, S. Weithoffer, C. d. Schryver, U. Wasenmüller, and N. Wehn. On parallel random number generation for accelerating simulations of communication systems. *Adv. Radio Sci.*, 12:75–81, 2014.
- [6] D. Chen, L. Wang, A. Zomaya, M. Dou, J. Chen, Z. Deng, and S. Hariri. Parallel simulation of complex evacuation scenarios with adaptive agent models. *IEEE Trans. Parall. Distr.*, 26(3):847–857, Mar. 2015.
- [7] S. Coakley, M. Gheorghe, M. Holcombe, S. Chin, D. Worth, and C. Greenough. Exploitation of high performance computing in the FLAME agent-based simulation framework. In *Proc. of the 14th Int. Conf. on High Performance Computing and Communications*, pages 538–545, Liverpool, UK, 2012. IEEE.
- [8] P. D. Coddington. Random number generators for parallel computers. Northeast Parallel Architecture Center, 1997. Paper 13.
- [9] N. Collier and M. North. Repast HPC: A platform for large-scale agent-based modeling. In W. Dubitzky, K. Kurowski, and B. Schott, editors, *Large-Scale Computing Techniques for Complex System Simulations*, chapter 5, pages 81–109. John Wiley & Sons, Inc., Dec. 2011.
- [10] B. Cosenza, G. Cordasco, R. De Chiara, and V. Scarano. Distributed load balancing for parallel agent-based simulations. In *Parallel, Distributed and Network-Based Processing (PDP), 2011 19th Euromicro International Conference on*, pages 62–69. IEEE, 2011.
- [11] R. D’Souza, M. Lysenko, S. Marino, and D. Kirschner. Data-parallel algorithms for agent-based model simulation of tuberculosis on graphics processing units. In *Proceedings of the 2009 Spring Simulation Multiconference*, SpringSim ’09, pages 21:1–21:12, San Diego, CA, USA, 2009. Society for Computer Simulation International.
- [12] R. D’Souza, M. Lysenko, and K. Rahmani. Sugarscape on steroids: simulating over a million agents at interactive rates. In *Proc. of Agent 2007 Conf.*, Chicago, USA, 2007.
- [13] D. W. Dyer. Uncommons Maths, 2012.
- [14] B. Edmonds and D. Hales. Replication, replication and replication: Some hard lessons from model alignment. *J. Artif. Soc. Soc. Simulat.*, 6(4):11, 2003.
- [15] N. Fachada. Agent-based simulation of the immune system. Master’s thesis, Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisboa, July 2008.
- [16] N. Fachada, V. V. Lopes, R. C. Martins, and A. C. Rosa. Towards a standard model for research in agent-based modeling and simulation. *PeerJ Comp. Sci.*, 1:e36, Nov. 2015.

- [17] N. Fachada, V. V. Lopes, and A. Rosa. Simulating antigenic drift and shift in influenza A. In *Proc. of the 2009 ACM symposium on Applied Computing*, SAC '09, pages 2093–2100, New York, NY, USA, 2009. ACM.
- [18] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design patterns: elements of reusable object-oriented software*. Pearson Education, 1994.
- [19] J. D. Gibbons and S. Chakraborti. *Nonparametric statistical inference*. Statistics: Textbooks & Monographs. Chapman and Hall/CRC, Boca Raton, FL, USA, fifth edition, July 2010.
- [20] K. Glass, M. Livingston, and J. Conery. Distributed simulation of spatially explicit ecological models. *SIGSIM Simul. Dig.*, 27(1):60–63, July 1997.
- [21] B. Göetz, T. Peierls, J. Bloch, J. Bowbeer, D. Holmes, and D. Lea. *Java concurrency in practice*. Addison-Wesley, 2006.
- [22] M. E. Goldsby and C. M. Pancerella. Multithreaded agent-based simulation. In *Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World*, WSC '13, pages 1581–1591, Washington, D.C., USA, 2013. IEEE Press.
- [23] Z. Gong, W. Tang, D. A. Bennett, and J.-C. Thill. Parallel agent-based simulation of individual-level spatial interactions within a multicore computing environment. *Int. J. Geogr. Inf. Sci.*, 27(6):1152–1170, 2013.
- [24] J. Gosling, B. Joy, G. Steele, G. Bracha, and A. Buckley. *The Java Language Specification (Java SE 8)*. Oracle Corporation, Redwood City, California, United States, 2014.
- [25] V. Grimm, U. Berger, D. DeAngelis, J. Polhill, J. Giske, and S. Railsback. The ODD protocol: A review and first update. *Ecol. Model.*, 221(23):2760–2768, 2010.
- [26] L. Gulyás, A. Szabó, R. Legéndi, T. Máhr, R. Bócsi, and G. Kampis. Tools for large scale (distributed) agent-based computational experiments. In *Proceedings of The Computational Social Science Society of the Americas*, 2011.
- [27] L. Gulyás, G. Szemes, G. Kampis, and W. de Back. A modeler-friendly API for ABM partitioning. In *Proceedings of the ASME 2009 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, pages 219–226, San Diego, CA, USA, Aug. 2009.
- [28] D. R. C. Hill, C. Mazel, J. Passerat-Palmbach, and M. K. Traore. Distribution of random streams for simulation practitioners. *Concurr. Comp.-Pract. Exp.*, 25(10):1427–1442, 2013.

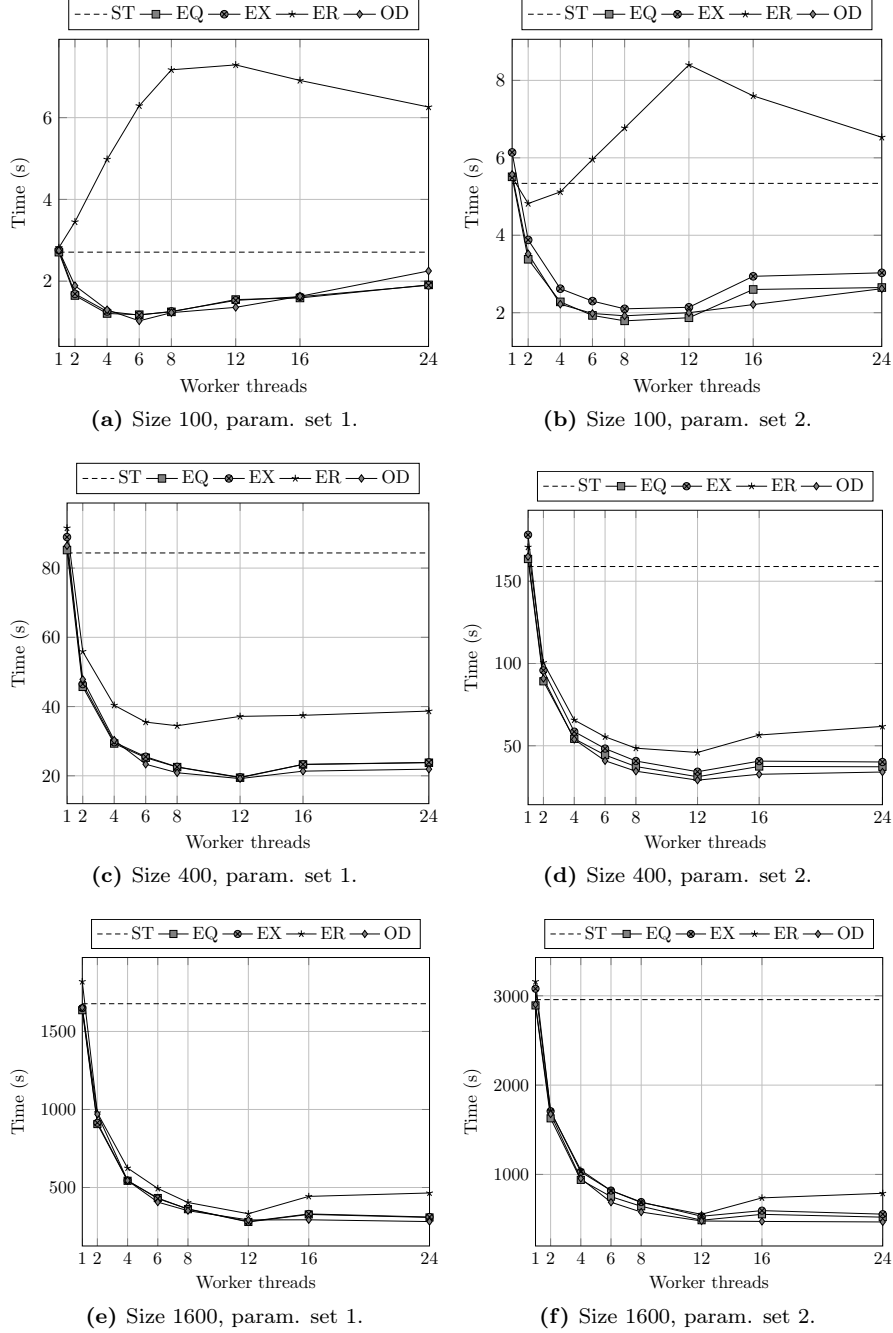
- [29] A. Husselmann and K. Hawick. Spatial agent-based modelling and simulations - a review. *Res. Lett. Inf. Math. Sci.*, 7:101–111, 2008.
- [30] A. G. Isaac. The ABM template models: A reformulation with reference implementations. *J. Artif. Soc. Soc. Simulat.*, 14(2):5, 2011.
- [31] M. Keenan, I. Komarov, R. M. D’Souza, and R. Riolo. Novel graphics processing unit-based parallel algorithms for understanding species diversity in forests. In *Proceedings of the 2012 Symposium on High Performance Computing*, page 10. Society for Computer Simulation International, 2012.
- [32] W. H. Kruskal and W. A. Wallis. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.*, 47(260):583–621, 1952.
- [33] A. M. Law. *Simulation Modeling and Analysis*. McGraw-Hill, Columbus, OH, USA, fifth edition, 2015.
- [34] D. E. Lenoski and W.-D. Weber. *Scalable shared-memory multiprocessing*. Elsevier, 1991.
- [35] S. Luke, C. Cioffi-Revilla, L. Panait, K. Sullivan, and G. Balan. MASON: A multiagent simulation environment. *Simulation*, 81(7):517–527, 2005.
- [36] M. Lysenko and R. D’Souza. A framework for megascale agent based model simulations on graphics processing units. *J. Artif. Soc. Soc. Simulat.*, 11(4):10, 2008.
- [37] S. L. Lytinen and S. F. Railsback. The evolution of agent-based simulation platforms: A review of NetLogo 5.0 and ReLogo. In *Proceedings of the Fourth International Symposium on Agent-Based Modeling and Simulation*, 2012.
- [38] C. Macal and M. North. Tutorial on agent-based modelling and simulation. *J. Simulat.*, 4(3):151–162, 2010.
- [39] U. Merlone, M. Sonnessa, and P. Terna. Horizontal and vertical multiple implementations in a model of industrial districts. *J. Artif. Soc. Soc. Simulat.*, 11(2):5, 2008.
- [40] B. Müller, S. Balbi, C. M. Buchmann, L. de Sousa, G. Dressler, J. Groeneveld, C. J. Klassert, Q. B. Le, J. D. Millington, H. Nolzen, D. C. Parker, J. G. Polhill, M. Schlüter, J. Schulze, N. Schwarz, Z. Sun, P. Taillandier, and H. Weise. Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environ. Model. Softw.*, 55:156–163, May 2014.
- [41] C. Nikolai and G. Madey. Tools of the trade: A survey of various agent based modeling platforms. *J. Artif. Soc. Soc. Simulat.*, 12(2):2, 2009.

- [42] M. North, N. Collier, and J. Vos. Experiences creating three implementations of the repast agent modeling toolkit. *ACM Trans. Model. Comput. Simul.*, 16(1):1–25, Jan. 2006.
- [43] M. J. North, N. T. Collier, J. Ozik, E. R. Tatara, C. M. Macal, M. Bragen, and P. Sydelko. Complex adaptive systems modeling with Repast Symphony. *Complex Adapt. Syst. Model.*, 1(1):1–26, 2013.
- [44] H. R. Parry and M. Bithell. Large scale agent-based modelling: A review and guidelines for model scaling. In A. J. Heppenstall, A. T. Crooks, L. M. See, and M. Batty, editors, *Agent-Based Models of Geographical Systems*, pages 271–308. Springer Netherlands, 2012.
- [45] K. S. Perumalla. Computational spectrum of agent model simulation. In S. Cakaj, editor, *Modeling Simulation and Optimization - Focus on Applications*, Numerical Analysis and Scientific Computing, chapter 12, pages 185–204. InTech, 2010.
- [46] S. Railsback, S. Lytinen, and S. Jackson. Agent-based simulation platforms: Review and development recommendations. *Simulation*, 82(9):609–623, 2006.
- [47] C. W. Reynolds. Flocks, herds and schools: A distributed behavioral model. *SIGGRAPH Comput. Graph.*, 21(4):25–34, July 1987.
- [48] P. Richmond, S. Coakley, and D. Romano. Cellular level agent based modelling on the graphics processing unit. In *High Performance Computational Systems Biology, 2009. HIBI’09. International Workshop on*, pages 43–50. IEEE, 2009.
- [49] P. Richmond, D. Walker, S. Coakley, and D. Romano. High performance cellular level agent-based simulation with FLAME for the GPU. *Brief. Bioinform.*, 11(3):334–347, 2010.
- [50] M. Scheutz and P. Schermerhorn. Adaptive algorithms for the dynamic distribution and parallel execution of agent-based models. *J. Parallel Distrib. Comput.*, 66(8):1037–1051, 2006.
- [51] E. Shook, S. Wang, and W. Tang. A communication-aware framework for parallel spatially explicit agent-based models. *Int. J. Geogr. Inf. Sci.*, 27(11):2160–2181, 2013.
- [52] A. Srinivasan, M. Mascagni, and D. Ceperley. Testing parallel random number generators. *Parallel Comput.*, 29(1):69–94, 2003.
- [53] G. L. Steele Jr, D. Lea, and C. H. Flood. Fast splittable pseudorandom number generators. In *OOPSLA ’14 Proceedings of the 2014 ACM International Conference on Object Oriented Programming Systems Languages & Applications*, pages 453–472. ACM, 2014.

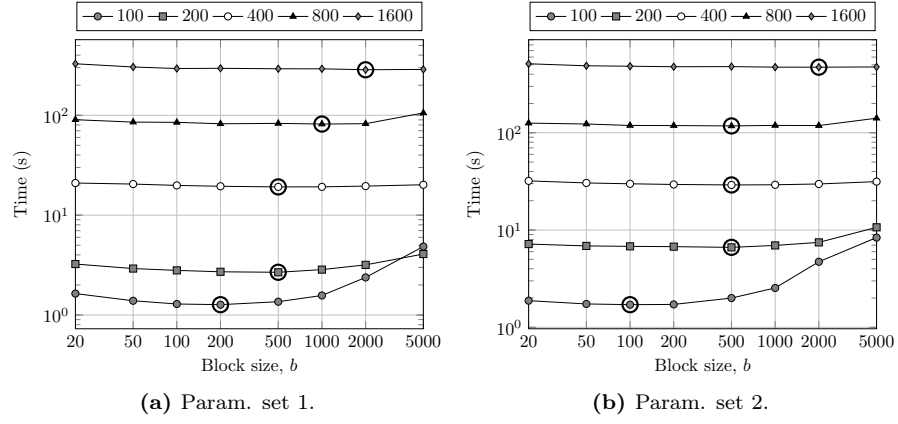
- [54] T. Takahashi and H. Mizuta. Efficient agent-based simulation framework for multi-node supercomputers. In *Proceedings of the 38th conference on Winter simulation*, pages 919–925. Winter Simulation Conference, 2006.
- [55] W. Tang and D. Bennett. Parallel agent-based modeling of spatial opinion diffusion accelerated using graphics processing units. *Ecol. Model.*, 222(19):3605–3615, 2011.
- [56] W. Tang and M. Jia. Global sensitivity analysis of a large agent-based model of spatial opinion exchange: A heterogeneous multi-GPU acceleration approach. *Ann. Assoc. Am. Geogr.*, 104(3):485–509, 2014.
- [57] W. Tang and S. Wang. HPABM: A hierarchical parallel simulation framework for spatially-explicit agent-based models. *Trans. GIS*, 13(3):315–333, 2009.
- [58] R. Tobias and C. Hofmann. Evaluation of free Java-libraries for social-scientific agent based simulation. *J. Artif. Soc. Soc. Simulat.*, 7(1):6, 2004.
- [59] L. G. Valiant. A bridging model for parallel computation. *Commun. ACM*, 33(8):103–111, Aug. 1990.
- [60] G. Viguera, J. M. Orduña, M. Lozano, J. M. Cecilia, and J. M. García. Accelerating collision detection for large-scale crowd simulation on multi-core and many-core architectures. *Int. J. High Perform. Comput. Appl.*, Feb. 2013.
- [61] A. Voss, J.-Y. You, E. Yen, H.-Y. Chen, S. Lin, A. Turner, and J.-P. Lin. Scalable social simulation: Investigating population-scale phenomena using commodity computing. In *e-Science (e-Science), 2010 IEEE Sixth International Conference on*, pages 1–8, Brisbane, QLD, Australia, Dec. 2010. IEEE.
- [62] J. Wang, N. Rubin, H. Wu, and S. Yalamanchili. Accelerating simulation of agent-based models on heterogeneous architectures. In *Proceedings of the 6th Workshop on General Purpose Processor Using Graphics Processing Units, GPGPU-6*, pages 108–119, Houston, Texas, USA, 2013. ACM.
- [63] K. Wang and Z. Shen. A GPU based trafficparallel simulation module of artificial transportation systems. In *Service Operations and Logistics, and Informatics (SOLI), 2012 IEEE International Conference on*, pages 160–165, Suzhou, China, 2012. IEEE.
- [64] G. Wei, P. Bogdan, and R. Marculescu. Efficient modeling and simulation of bacteria-based nanonetworks with BNSim. *IEEE J. Select. Areas Commun.*, 31(12):868–878, Dec. 2013.
- [65] U. Wilensky. NetLogo, 1999.
- [66] U. Wilensky. NetLogo heatbugs model, 2004.



- [67] U. Wilensky and W. Rand. Making models match: replicating an agent-based model. *J. Artif. Soc. Soc. Simulat.*, 10(4):2, 2007.
- [68] X.-J. Yang, X.-K. Liao, K. Lu, Q.-F. Hu, J.-Q. Song, and J.-S. Su. The TianHe-1A supercomputer: its hardware and software. *J. Comput. Sci. Tech.*, 26(3):344–351, 2011.
- [69] M. Yokokawa, F. Shoji, A. Uno, M. Kurokawa, and T. Watanabe. The K computer: Japanese next-generation supercomputer development project. In *Proceedings of the 17th IEEE/ACM international symposium on Low-power electronics and design*, pages 371–372. IEEE, 2011.



**Figure 6** – Simulation time versus number of workers for both parameter sets and model sizes 100, 400 and 1600, with  $b = 500$  for the OD variant. Time for single-thread (ST) implementation shown as reference.



**Figure 7** – OD performance with 12 workers for the two parameter sets and different model sizes. Best performance marked with bold circle.

Version	Size	Param. set 1				Param. set 2			
		$\bar{t}(s)$	$s(\%)$	$S_p^{NL}$	$S_p^{ST}$	$\bar{t}(s)$	$s(\%)$	$S_p^{NL}$	$S_p^{ST}$
NL	100	15.86	2.26	1.00	0.17	32.18	2.13	1.00	0.17
	200	100.25	1.25	1.00	0.12	245.38	0.61	1.00	0.15
	400	481.48	1.25	1.00	0.18	1074.21	0.34	1.00	0.15
	800	2077.10	0.47	1.00	0.18	4536.90	0.51	1.00	0.15
	1600	9115.80	1.03	1.00	0.18	19559.30	0.46	1.00	0.15
ST	100	2.71	0.82	5.85	1.00	5.34	0.96	6.03	1.00
	200	12.17	1.80	8.24	1.00	36.12	0.49	6.79	1.00
	400	84.37	3.35	5.71	1.00	158.95	0.30	6.76	1.00
	800	382.63	1.32	5.43	1.00	699.59	0.52	6.49	1.00
	1600	1677.82	4.67	5.43	1.00	2957.20	4.15	6.61	1.00
EQ	100	1.55	1.62	10.24	1.75	1.87	1.53	17.21	2.85
	200	2.81	4.01	35.61	4.32	7.08	1.78	34.64	5.10
	400	19.46	1.10	24.74	4.34	31.17	0.66	34.46	5.10
	800	86.08	4.95	24.13	4.45	125.27	3.32	36.22	5.58
	1600	279.23	1.45	32.65	6.01	487.34	1.74	40.14	6.07
EX	100	1.53	1.90	10.39	1.78	2.14	2.75	15.06	2.50
	200	2.91	3.69	34.40	4.18	8.08	1.74	30.37	4.47
	400	19.56	1.54	24.62	4.31	34.22	1.54	31.40	4.65
	800	86.49	6.31	24.01	4.42	138.99	4.29	32.64	5.03
	1600	281.57	1.95	32.37	5.96	531.96	0.99	36.77	5.56
ER	100	7.29	4.46	2.18	0.37	8.39	1.76	3.83	0.64
	200	16.44	4.68	6.10	0.74	17.91	1.41	13.70	2.02
	400	37.16	0.55	12.96	2.27	45.91	0.62	23.40	3.46
	800	111.45	3.02	18.64	3.43	159.25	2.02	28.49	4.39
	1600	331.77	1.06	27.48	5.06	553.44	1.45	35.34	5.34
OD	100	1.36	1.16	11.70	2.00	2.00	1.66	16.13	2.68
	200	2.68	2.61	37.42	4.54	6.64	1.64	36.95	5.44
	400	19.19	1.04	25.09	4.40	29.09	0.42	36.93	5.46
	800	82.94	2.73	25.04	4.61	117.62	2.55	38.57	5.95
	1600	292.16	2.91	31.20	5.74	478.83	1.95	40.85	6.18

**Table 7** – Times and speedups for the different versions using both parameter sets and tested model sizes. The  $\bar{t}(s)$  column specifies the mean simulation time for each version and model size combination. The  $s(\%)$  column shows the associated relative standard deviation, given by  $100 \cdot s/\bar{t}(s)$ , where  $s$  is the sample standard deviation. The  $S_p^{NL}$  and  $S_p^{ST}$  columns display the speedups verified against the NetLogo and Java single-thread versions, respectively. The number of workers,  $p$ , is set to 12 for the parallel variants, and 1 for the NL and ST versions. For the OD variant,  $b = 500$ .

Output	Stat.	Param. set 1					Param. set 2				
		100	200	400	800	1600	100	200	400	800	1600
$P_i^s$	max	0.777	<u>0.046</u>	0.283	0.833	0.832	<u>0.018</u>	0.203	0.894	0.584	0.418
	arg max	0.785	0.456	0.990	<u>0.030</u>	0.575	0.685	0.600	0.785	0.445	1.000
	min	0.776	<u>0.029</u>	0.086	<u><u>0.007</u></u>	0.161	0.191	0.323	0.539	1.000	1.000
	arg min	0.546	0.679	0.542	0.085	0.348	0.555	0.169	0.539	1.000	1.000
	$\bar{X}^{ss}$	0.489	0.984	0.298	0.533	0.976	0.683	0.079	0.204	0.886	0.199
	$S^{ss}$	0.054	0.610	0.572	0.790	<u>0.025</u>	0.729	0.194	0.530	0.387	0.159
$P_i^w$	max	0.062	0.981	0.449	0.484	0.685	0.436	0.533	0.205	0.249	<u>0.008</u>
	arg max	0.944	0.494	0.658	0.579	0.469	0.443	0.667	0.243	0.793	0.698
	min	0.615	0.168	0.447	0.675	0.438	0.081	0.720	0.483	0.111	0.483
	arg min	0.381	0.407	0.382	0.410	0.630	0.846	0.444	0.084	0.362	0.707
	$\bar{X}^{ss}$	0.428	0.995	0.815	0.729	0.499	0.785	0.177	0.893	0.632	0.636
	$S^{ss}$	0.161	0.795	0.840	0.604	<u>0.003</u>	0.700	0.268	0.667	0.350	0.185
$P_i^c$	max	0.717	0.565	0.144	0.054	<u>0.025</u>	0.298	0.542	0.129	0.583	0.036
	arg max	0.383	0.749	0.582	0.416	1.000	<u>0.050</u>	0.534	0.235	0.739	0.463
	min	0.766	0.122	0.593	0.379	0.854	0.118	0.209	0.644	0.426	0.217
	arg min	0.942	0.802	0.067	0.896	0.823	0.822	0.862	0.413	0.302	0.671
	$\bar{X}^{ss}$	0.566	0.970	0.386	0.488	0.887	0.956	0.453	0.067	0.421	0.367
	$S^{ss}$	0.073	0.543	0.644	0.785	<u>0.011</u>	0.723	0.255	0.644	0.358	0.123
$\bar{E}_i^s$	max	0.595	0.183	0.410	0.797	0.718	0.489	0.111	0.538	<u>0.004</u>	0.115
	arg max	0.970	0.169	0.443	0.180	0.369	0.940	0.520	0.819	0.348	1.000
	min	<u>0.003</u>	<u>2e-5</u>	0.142	<u>4e-6</u>	0.099	0.400	0.077	0.967	0.102	0.012
	arg min	1.000	1.000	1.000	1.000	1.000	0.705	0.485	0.871	0.566	<u>0.627</u>
	$\bar{X}^{ss}$	0.901	0.236	0.909	0.981	0.987	0.451	<u>0.024</u>	0.276	0.372	0.216
	$S^{ss}$	0.242	0.633	0.200	0.480	0.250	0.642	0.458	0.395	0.425	0.250
$\bar{E}_i^w$	max	0.508	0.664	0.696	0.672	0.716	0.269	0.101	0.619	0.944	0.331
	arg max	0.338	0.687	0.156	0.247	0.467	0.499	0.620	0.057	0.974	0.371
	min	0.518	<u>0.019</u>	<u>0.005</u>	0.143	<u>0.020</u>	0.074	0.973	0.079	0.145	0.591
	arg min	0.109	0.963	0.514	0.425	0.278	0.598	<u>0.033</u>	0.459	0.317	0.949
	$\bar{X}^{ss}$	<u>0.026</u>	0.320	0.224	0.204	0.686	0.971	<u>0.010</u>	0.345	0.876	0.212
	$S^{ss}$	0.547	0.897	0.741	0.986	0.061	0.663	0.400	0.665	0.457	0.115
$\bar{C}_i$	max	0.795	0.129	0.622	0.313	0.742	<u>0.045</u>	0.823	0.270	0.717	0.118
	arg max	0.674	0.792	<u>0.044</u>	0.996	0.772	<u>0.257</u>	0.581	0.234	0.735	0.276
	min	0.612	0.712	<u>0.164</u>	0.118	0.056	0.261	0.586	0.165	0.640	<u>0.037</u>
	arg min	0.709	0.834	1.000	1.000	1.000	0.078	0.570	0.547	0.901	0.904
	$\bar{X}^{ss}$	0.585	0.975	0.365	0.509	0.880	0.966	0.423	0.072	0.404	0.326
	$S^{ss}$	0.073	0.539	0.644	0.776	<u>0.010</u>	0.702	0.250	0.647	0.367	0.126

**Table 8** –  $P$ -values for the Kruskal-Wallis statistical test which tests the null hypothesis that the specified focal measures from the six PPHPC versions are drawn from the same distribution. Smaller  $p$ -values (lower than 0.05 or 0.01) question the null hypothesis, indicating that at least one implementation produced an output significantly different from the others. Values lower than 0.05 are underlined, while values lower than 0.01 are double-underlined. The number of workers,  $p$ , is set to 12 for the parallel implementations. For the OD implementation,  $b = 500$ .